

What Is the Impact of Additional Tropical Observations on a Modern Data Assimilation System?

LAURA C. SLIVINSKI AND GILBERT P. COMPO

*Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and
Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado*

JEFFREY S. WHITAKER

Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado

PRASHANT D. SARDESHMUKH AND JIH-WANG A. WANG

*Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and
Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado*

KATE FRIEDMAN

NOAA/NWS/NCEP/Environmental Modeling Center, College Park, Maryland

CHESLEY MCCOLL

*Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, and
Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado*

(Manuscript received 30 March 2018, in final form 30 April 2019)

ABSTRACT

Given the network of satellite and aircraft observations around the globe, do additional in situ observations impact analyses within a global forecast system? Despite the dense observational network at many levels in the tropical troposphere, assimilating additional sounding observations taken in the eastern tropical Pacific Ocean during the 2016 El Niño Rapid Response (ENRR) locally improves wind, temperature, and humidity 6-h forecasts using a modern assimilation system. Fields from a 50-km reanalysis that assimilates all available observations, including those taken during the ENRR, are compared with those from an otherwise-identical reanalysis that denies all ENRR observations. These observations reveal a bias in the 200-hPa divergence of the assimilating model during a strong El Niño. While the existing observational network partially corrects this bias, the ENRR observations provide a stronger mean correction in the analysis. Significant improvements in the mean-square fit of the first-guess fields to the assimilated ENRR observations demonstrate that they are valuable within the existing network. The effects of the ENRR observations are pronounced in levels of the troposphere that are sparsely observed, particularly 500–800 hPa. Assimilating ENRR observations has mixed effects on the mean-square difference with nearby non-ENRR observations. Using a similar system but with a higher-resolution forecast model yields comparable results to the lower-resolution system. These findings imply a limited improvement in large-scale forecast variability from additional in situ observations, but significant improvements in local 6-h forecasts.

1. Introduction

The El Niño–Southern Oscillation (ENSO) is a coupled ocean–atmosphere phenomenon that has global

climatic effects in addition to the regional effects on the tropical Pacific Ocean, where it originates (Ropelewski and Halpert 1987; Halpert and Ropelewski 1992; Neelin et al. 1998; Trenberth et al. 1998; Barsugli and Sardeshmukh 2002; McPhaden et al. 2006; Sardeshmukh et al. 2000; Compo et al. 2001). Understanding and forecasting atmospheric anomalies associated with

Corresponding author: Laura C. Slivinski, laura.slivinski@noaa.gov

DOI: 10.1175/MWR-D-18-0120.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](https://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

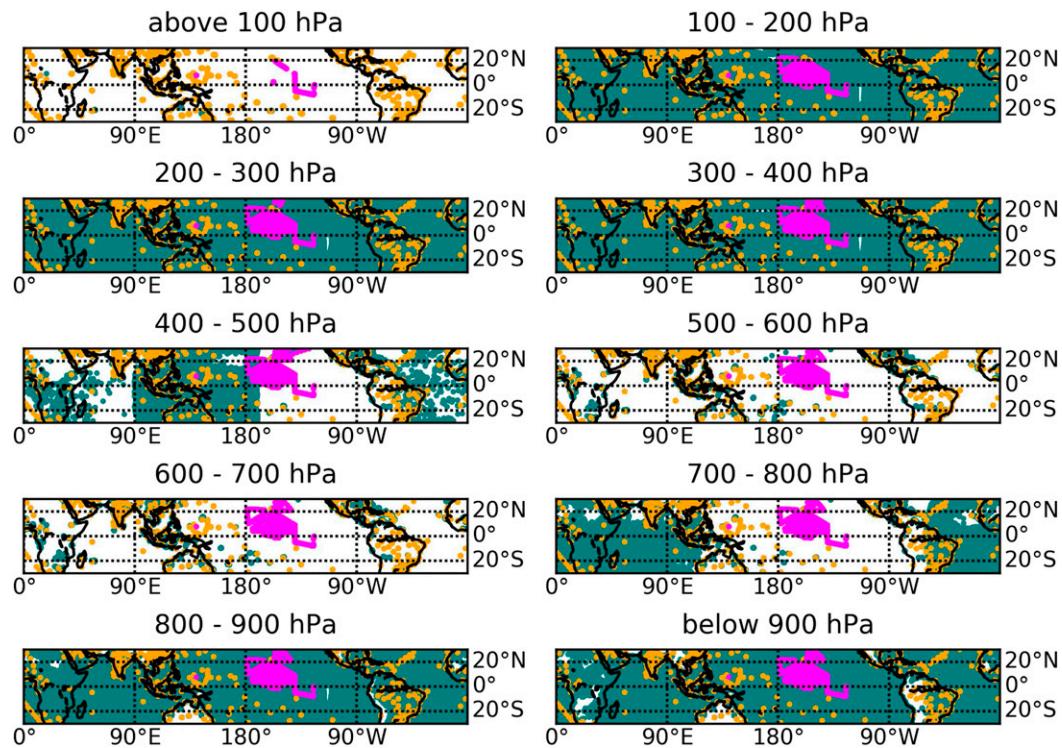


FIG. 1. Maps of meridional wind observations assimilated into the NCEP operational GFS between 22 Jan and 7 Mar 2016 at the indicated levels (30°S–30°N only). Orange points represent non-ENRR radiosondes, and magenta shows ENRR observations. Teal points represent all remaining observations, including aircraft observations and AMVs.

ENSO are of particular importance as these can be high-impact and extreme weather and climate events (Kiladis and Diaz 1989; Glantz 2001). While onboard observation campaigns have targeted the tropical Pacific Ocean during El Niños (Kashino et al. 2009), prior to 2015, no field campaigns had been organized to take atmospheric observations in the tropical Pacific during an El Niño. By mid-2015, forecasts predicted a strong El Niño for the 2015/16 winter (L’Heureux et al. 2017), providing the National Oceanic and Atmospheric Administration (NOAA) the opportunity to initiate a campaign to collect observations in the tropical Pacific Ocean during this event. NOAA’s El Niño Rapid Response (ENRR) project undertook the first field campaign to take atmospheric observations in this region during an El Niño (Dole et al. 2018).

Although the existing wind observation network is dense in the tropics near the surface and in the upper troposphere, there is a dearth of observations at other levels, particularly over the Pacific Ocean (orange and teal dots in Fig. 1). While aircraft measure winds and satellites provide them via atmospheric motion vectors (AMVs; Velden et al. 1997), the in situ observation networks of humidity (Fig. 2) and temperature (not

shown) are even sparser. The ENRR campaign sought to fill this void and observe the strong El Niño as it was under way. The campaign was successfully executed and many observations were taken from land, sea, and air platforms in early 2016; see magenta dots in Figs. 1 and 2 (Dole et al. 2018; Hartten et al. 2018a, b). These observations were transmitted to the Global Telecommunications System and assimilated into operational weather forecasting systems.

A summary of wind observations taken in the ENRR campaign period from 22 January to 7 March 2016 is provided in Table 1. Region A is shown in Fig. 3. Note that there are more ENRR observations than observations taken in Region A because the ship and aircraft released radiosondes and dropsondes, respectively, along their routes from the United States to the tropical Pacific Ocean.

To systematically study the effect of these observations on analyses and short-term forecasts from the Global Forecast System (GFS) of the NOAA/National Centers for Environmental Prediction (NCEP), a suite of “reanalysis” datasets was generated for the 3 months spanning the ENRR campaign period. The “control” reanalysis assimilated all observations, including the

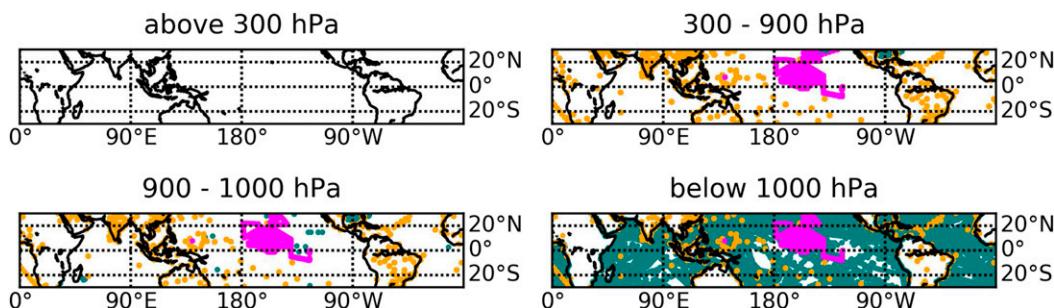


FIG. 2. Maps of in situ specific humidity observations assimilated into the NCEP operational weather forecast system between 22 Jan and 7 Mar 2016 at the indicated levels (30°S–30°N only). Orange points represent non-ENRR radiosondes and magenta shows ENRR observations. Teal points represent all remaining observations.

ENRR observations, while the “denial” reanalysis was generated identically to, and in parallel to, the control experiment except that it never assimilated any ENRR observations.

Data-denial experiments, or “observing system experiments,” are commonly used to determine observation impact, by either removing part of the current observational network from a reanalysis or by adding targeted observations from field campaigns (Kelly et al. 2007; Gelaro and Zhu 2009). Most studies are limited to the extratropics (Bouttier and Kelly 2001; Benjamin et al. 2004; Anwender et al. 2012; Hamill et al. 2013; Majumdar et al. 2013; Romine et al. 2016), and those that target the tropics are generally focused on forecasting tropical cyclones and typhoons (Harnisch and Weissmann 2010; Torn 2014). Langland (2005) and Majumdar (2016) provide reviews of recent targeted observation campaigns, including during the World Meteorological Organization/World Weather Research Programme’s The Observing System Research and Predictability Experiment (THORPEX; Gelaro et al. 2010), and the associated observation impact studies. In general, the influence of targeted observations is positive but often very small, particularly with modern data assimilation systems and dense observational networks.

TABLE 1. Total approximate number of wind observations (in thousands of observations) throughout the atmosphere over the 6-week research period 22 Jan–7 Mar 2016.

Platform	Global	Region A	ENRR
Rawinsonde	3358	5.4	8.5
Dropsonde	22.5	17	21
Pibal	189	0	0
Wind profiler	41.4	0	0
Aircraft	22 039	9.8	0
Satellite wind	27 606	2255	0
Surface marine	1143	20.5	0
Scatterometer	1553	62.3	0
Total	55 951.9	2370	29.5

As there have been no other atmospheric field campaigns targeting the deep tropics during an El Niño, it is not known whether these past results apply to such a situation.

Here, two sets of data-denial experiments at different resolutions were conducted to determine the impact of the ENRR observations. Results from these experiments demonstrate that the ENRR observations have a significant impact on the analysis fields, particularly meridional wind at 200 hPa, suggesting that the existing observational network is not dense enough in this region to make the ENRR observations redundant. The 6-hourly background fields were also significantly improved when ENRR observations were assimilated. The results further suggest that the additional observations had small effects on the assimilation of nearby AMVs, and almost no effect on the fit to other nearby non-ENRR observations.

The paper is organized as follows. Section 2 describes the setup for the data-denial experiments, including an overview of the observations during the ENRR campaign. Section 3 describes the regional effects of the ENRR observations on GFS analyses of 200-hPa

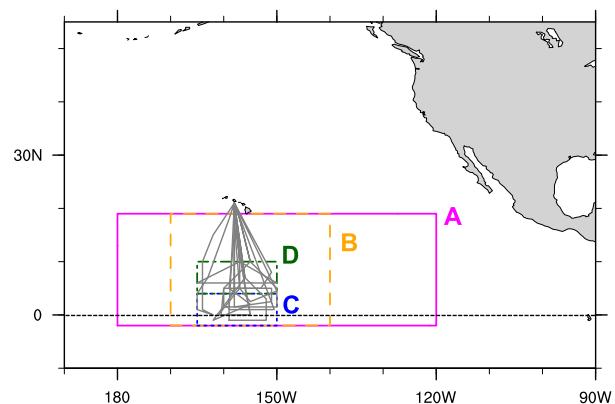


FIG. 3. Map of “deep tropics” flights (gray lines) with regions labeled A–D; see text for details.

meridional wind. Section 4 describes local effects of these observations on short-term forecasts and analyses of wind, humidity, and temperature. Comparisons to non-ENRR observations, including in situ observations and measurements from satellites and aircraft, are given in section 5. A summary of the same performance metrics on the second set of experiments with a higher-resolution model is discussed in section 6. Finally, section 7 includes discussion and conclusions.

2. Data-denial experiments

Two sets of two experiments were completed: low-resolution (“low-res”) experiments with and without the ENRR sounding observations assimilated, and high-resolution (“high-res”) experiments with and without the ENRR sounding observations (hereafter, ENRR observations). All experiments used the NCEP GFS version Q3FY16 and assimilated observations via the hybrid four-dimensional ensemble variational (4DEnVar) algorithm (Kleist and Ide 2015) with an ensemble of 80 members. The low-res experiments were conducted at a resolution of T254, or about 50 km, and the high-res experiments utilized an ensemble at resolution T574 (about 34 km) with a separate control member at a resolution of T1534 (about 13 km). In all cases, the vertical resolution consisted of 64 vertical levels. Note that the resolution of the GFS that was operational during the ENRR campaign is the same as that of the high-res experiments (T574 ensemble/T1534 control), though the version of the GFS used in the experiments below is newer than the version that was operational during the campaign.

The 4DEnVar algorithm consists of forecasting an ensemble of model states (the “background” or “first-guess” ensemble), in addition to a control member. The control member is then updated using the hybrid gridpoint statistical interpolation (GSI) analysis scheme (Hu et al. 2016; Wang et al. 2013), which uses a weighted combination of the ensemble background covariance and a static background covariance. In this implementation, the ensemble is updated through an ensemble Kalman filter (EnKF) analysis step, and then recentered around the control member, resulting in the “analysis” ensemble.

After the analysis is calculated, a smoothing treatment is applied to prevent gravity wave noise from dominating the short-term evolution of the forecast following this update. Note that the low-res and high-res experiments use different treatments. The low-res experiments use a 4D incremental analysis update (IAU; Bloom et al. 1996; Lei and Whitaker 2016), not implemented operationally at the time of the experiments. In the IAU,

observations are assimilated in a 6-h time window to calculate an “initial” analysis, as usual; instead of forecasting from these initial analysis fields, though, the “final” analysis fields are obtained by applying the initial analysis increments as a forcing to the background ensemble in intermediate steps. In other words, instead of applying the increments as a discrete jump at each analysis time, they are interpolated in time and applied within the model, between assimilation times. Note that, for the low-res results included here, the increments shown are from the initial analysis fields, before IAU was applied; results are qualitatively similar but weaker for the final fields, after the IAU is applied (see the appendix). Finally, in the version of the 4DEnVar used for the low-res experiments, the control member is set equal to the ensemble mean, so there is no “recentering” of the ensemble around a separate control member. In contrast, the operational 4DEnVar, and the high-res experiments described below, use a separate, higher-resolution control member.

The high-res experiments used the version of the hybrid 4DEnVar that became operational on 11 May 2016. This is very similar to the algorithm used in the low-res experiments described above, but without 4DIAU. Instead, the NCEP operational digital filter was applied to the background fields to prevent numerical instabilities from forming (Lynch and Huang 1992; Huang and Lynch 1993).

ENRR observation platforms include dropsondes (Vaisala RS-92) from 22 flights of the NOAA Gulfstream-IV aircraft (hereafter, G-IV) and 6 coordinated flights of Air Force C130s (NOAA 2018; UCAR/NCAR–Earth Observing Laboratory 1994). ENRR observations also include radiosondes (Vaisala RS92-SGP) launched from the NOAA ship *Ronald H. Brown* (RHB; Cox et al. 2017) and a field campaign station on Kiritimati Island (Dole et al. 2018; Hartten et al. 2018a, b, 2017). (Observation data from this campaign are available at https://www.esrl.noaa.gov/psd/enso/rapid_response/data_pub/.) The aircraft component of the campaign began with a G-IV flight on 22 January 2016, and concluded on 7 March 2016. Ten of these flights are designated “deep tropics flights” that flew into the region 2°S–10°N, 150°–165°W. See Fig. 3 for a map of the approximate flight paths for these 10 flights, as well as regions A–D considered in section 5. The deep tropics flight dates are 21, 25, and 26 January; 2, 12, 15, 26, 27, and 29 February; and 1 March 2016. Note that this subset is similar, but not identical to, the “convective enclosure” flights defined in Dole et al. (2018); see Fig. 1 of that work. Prior investigations (not shown) suggest that results shown here (Figs. 5 and 14) would not change significantly using the

convective enclosure flights. Observations from these flights were generally assimilated at 0000 UTC on the day following the local flight date; that is, most observations taken on the 21 January flight were assimilated at 0000 UTC 22 January. The *Ronald H. Brown* sailed from 17 February to 17 March 2016, launching radiosondes every 3–4 h. Radiosondes were also launched from Kiritimati Island every 12 h; they were assimilated from 16 February 2016 through the end of the campaign on 31 March 2016. Magenta points on Fig. 1 represent all ENRR observations of wind within the period from 14 February through 7 March 2016; this includes all radiosondes launched from Kiritimati Island and the *Ronald H. Brown* and all dropsondes released from all ENRR flights in that period. These instruments took measurements of pressure, wind, temperature, and humidity. See Dole et al. (2018) for a complete description of the ENRR observations.

To rigorously determine the effects of the ENRR observations on analyses and short-term forecasts within the data assimilation system, two experiments were run in each of the setups described above: first, a retrospective analysis covering the entire campaign period (20 January–31 March 2016) that assimilates all observations considered in the real-time operational forecasts, including the ENRR observations (the control experiment); and second, an experiment identical to the control but that never assimilates the ENRR observations (the denial experiment), run in parallel to the control. Surface-level observations from the ENRR platforms were not denied, since the *Ronald H. Brown* and Kiritimati Island take surface observations in normal (non-ENRR campaign) conditions. These surface-level observations account for about 0.3% of the total number of ENRR observations. Any differences between the control and denial experiments should theoretically be due solely to the assimilation of the ENRR observations. In practice though, the model includes a version of the stochastically perturbed parameterization tendencies scheme with perturbed boundary layer humidity (SPPT/SHUM; Palmer et al. 2009; Tompkins and Berner 2008); see Wang et al. (2019) for a description of the implementation of the scheme in this model. Thus, some differences between the control and denial experiments will be due to different realizations of the stochastic perturbations. The spectral forecast model may create nonlocal differences between two realizations as well.

Note that the impact of additional observations within a data assimilation system will depend on the uncertainties in the background field and in the observations themselves. If the background uncertainty is relatively low and the error of the additional

observations is relatively high, then the assimilation system will effectively ignore the observations. In contrast, if there is greater background uncertainty and the new observations are relatively accurate, then even a small amount of new observations can make an impact. In the experiments examined here, the observation errors and background uncertainty both vary significantly in time and space. Observation errors depend on platform, variable observed, and vertical level (Hu et al. 2016; Bormann et al. 2003). For example, dropsonde wind errors vary from 2.4 m s^{-1} at the surface, increasing to 3.4 m s^{-1} at 300 hPa, and decreasing again to 2.7 m s^{-1} at 50 hPa and above. Satellite wind errors also vary among satellites themselves: EUMETSAT observations from Meteosat have errors prescribed as 3.8 m s^{-1} at the surface, increasing to 7 m s^{-1} at 250 hPa and above; NESDIS observations from GOES are prescribed as 7.6 m s^{-1} at the surface, increasing to 14 m s^{-1} at 250 hPa and above. Additionally, these experiments use a hybrid parameter value of $\beta = 0.125$, where $\beta = 0$ implies a background covariance determined entirely by the ensemble, and $\beta = 1$ yields an entirely static background covariance. Therefore, the background uncertainty is mostly determined by the prior ensemble covariance, and so it will also vary based on variable, spatiotemporal location, and dynamical situation. (See Kleist and Ide (2015) for a discussion of how the hybrid increments depend on the covariances.) As an example, the first-guess ensemble spread at Kiritimati Island for meridional wind during the campaign period varies from 0.5 to 2 m s^{-1} at surface levels, and from 2 to 5 m s^{-1} at 200 hPa and above; these numbers depend on synoptic conditions and time of day as well. The results from the data-denial experiments studied here will demonstrate whether the assimilation system deemed the ENRR observations accurate and useful enough, relative to the background uncertainty, to make an impact within the existing observation network. In all quantities considered here, the results are tested against the null hypothesis that they do not make an impact.

3. Regional effects of observations

El Niño is generally characterized by increased divergence in the upper troposphere above the eastern equatorial Pacific Ocean, which should manifest itself in the 200-hPa meridional wind field (e.g., Trenberth et al. 1998). Upper-level divergent outflow is critical for generating El Niño's global extratropical-tropical teleconnections (e.g., Sardeshmukh and Hoskins 1988). For this feature, the operational GFS forecast was systematically weaker than observations in this time period

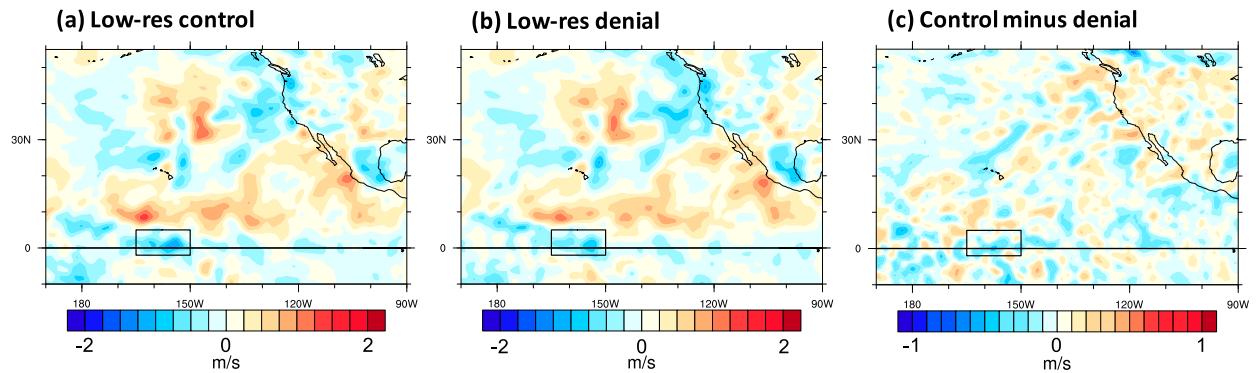


FIG. 4. Assimilation increments of 200-hPa meridional wind averaged over 22 Jan–7 Mar, valid at 0000 UTC: (a) low-res control experiment, (b) denial experiment, and (c) the difference. The black box emphasizes the deep tropics region (see text for details).

[Fig. 11 of Dole et al. (2018)]. Thus, when observations targeted to this area are assimilated, we expect to see greater divergence aloft if the newer version of the GFS employed here has the same deficiencies as the version of the GFS that was operational during the ENRR campaign. Figures 4 and 5 show the time-averaged increments (differences between the analysis and background) of the 200-hPa meridional wind field. In both the control and denial experiments over the 22 January–7 March 2016 time period (Fig. 4), there is a large-scale pattern of outflow north of the equator (particularly in the region 2°S–15°N, 150°–170°W). Figure 4c shows the difference between the average control and denial increments over this period. Recall that these experiments use a spectral model with stochastic perturbations; this will yield nonzero differences between the control and denial increments that are unrelated to the ENRR observations. To determine whether the differences between the control and denial increments are significant, a Kolmogorov–Smirnov (DKS) test (Massey 1951) is calculated, and

cross hatching is included on the control-denial difference figures where the significance is at least 95%. Comparisons between the control experiment and the denial experiment for 22 January–7 March 2016 suggest that assimilating ENRR observations may lead to stronger outflow from the deep tropics area than when ENRR observations are not assimilated, but the differences between the control and denial fields over this period are small and not statistically significant at the 95% level (as shown by the lack of cross hatching on Fig. 4c).

The effect of assimilating ENRR observations on the strengthening of the outflow is more apparent and significant on days when flights entered the deep tropics area: Fig. 5 is similar to Fig. 4, but restricted to “deep tropics” flight days. The difference map (Fig. 5c) shows the strongest signals in the deep tropics box. This difference is not significant at the 95% level according to a DKS test (see the lack of cross hatching on Fig. 5c), but there are several grid points in the deep tropics box for which it is statistically significant at the 90% level

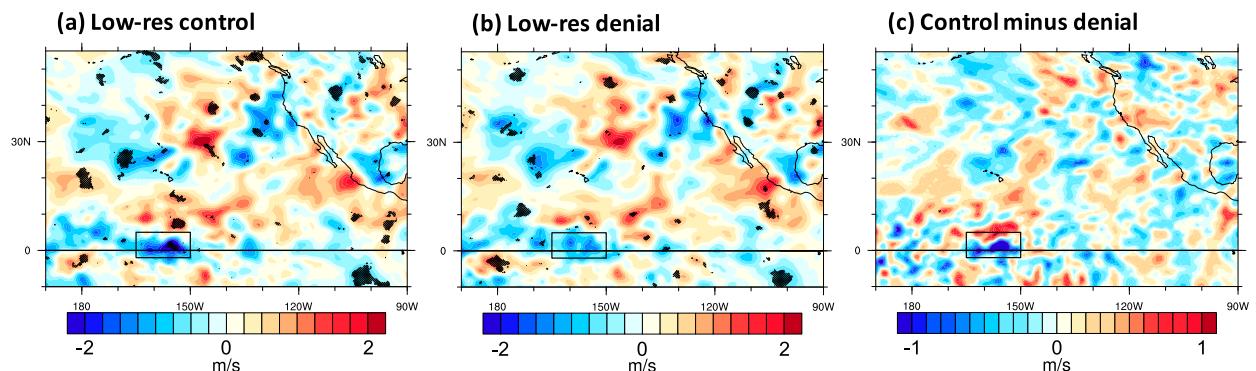


FIG. 5. Assimilation increments of 200-hPa meridional wind averaged over days with flights into the “deep tropics,” valid at 0000 UTC: (a) low-res control experiment, (b) denial experiment, and (c) the difference between the two. In (a),(b), cross hatching represents significant differences from respective nonflight days (95% level) based on a Kolmogorov–Smirnov test. In (c), the lack of cross hatching demonstrates the lack of significant differences from zero at the 95% level. The black box emphasizes the deep tropics region.

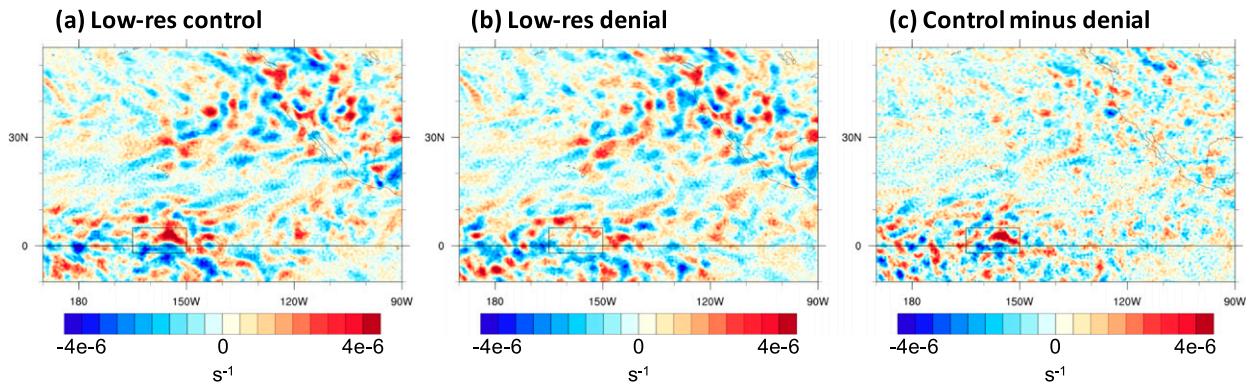


FIG. 6. Assimilation increments of 200-hPa divergence averaged over days with flights into the “deep tropics,” valid at 0000 UTC: (a) low-res control experiment, (b) denial experiment, and (c) the difference between the two. The black box emphasizes the deep tropics region.

(not shown). Since this is based on only 10 pairs of data, the significance level is not very strong. The 200-hPa divergence increments (Fig. 6) show that the stronger 200-hPa outflow is indeed associated with upper-level divergence in the deep tropics area. Several of the largest positive values (dark red) in the deep tropics box of Fig. 6c are significant at the 90% level, with some grid points above the 95% level.

As previously stated, these figures are calculated from the initial analysis fields, before the IAU is applied in that time window. When the final (post-IAU forcing) analysis fields are used, the results are qualitatively similar, but the magnitudes of the increments are diminished by about 60%–80% (Fig. A1). This suggests that the final analysis fields are closer to the forecast fields than the initial analysis used to calculate the IAU forcing. In principle, this could imply a smaller impact of any observation on the final analysis when IAU is utilized. Deeper investigations into the IAU algorithm and its effects on data assimilation increments, particularly in the context of individual weather events, will be explored in future work.

To investigate whether the stronger outflow in the upper level, which indicates stronger convergence below, is associated with improved estimates of precipitation, the control and denial fields are compared with precipitation estimates from the NASA Global Precipitation Measurement (GPM) *Core Observatory* satellite data. These data used the Tropical Rainfall Measuring Mission (TRMM) 3B42 algorithm, and are provided at a resolution of 0.1°. Figure 7 illustrates the RMS differences between the precipitation rates from the given experiment and from the GPM (both interpolated to a 0.25° grid); the mean is taken in time for 0000–0300 UTC on deep tropics flight days. Figure 7c shows the difference field (denial RMSD subtracted from the control RMSD); the lack of cross hatching indicates that there were no regions with a statistically significant pattern above the 80% level. The spatial average of the control (denial) RMSD over regions C and D (Fig. 3) is 0.94 mm h⁻¹ (0.97 mm h⁻¹). While not significant, the control RMSD is smaller than the denial RMSD. The consistent increment of the analyzed upper-level outflow relative to the background fields by the

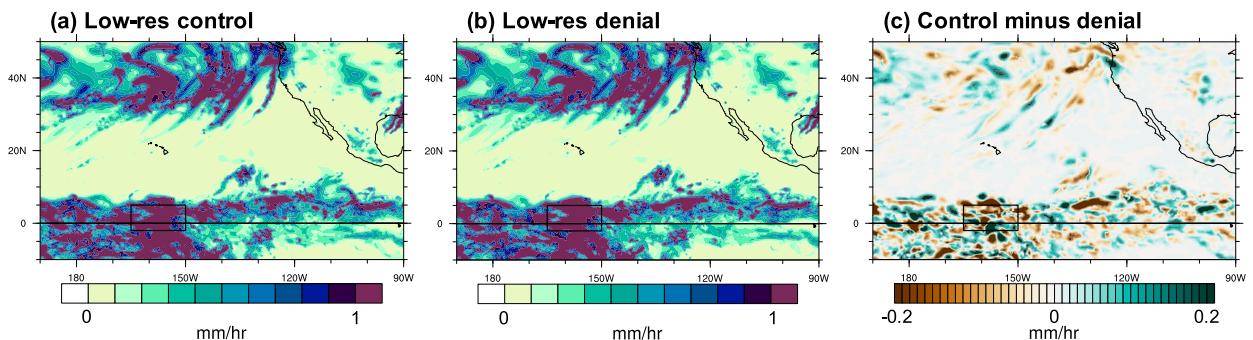


FIG. 7. Precipitation rate RMS differences from the NASA GPM dataset for (a) the low-res control and (b) denial experiments, averaged over 0000–0300 UTC on deep tropics flight days. (c) The difference between (a) and (b); note the different color scales. The black box emphasizes the deep tropics region.

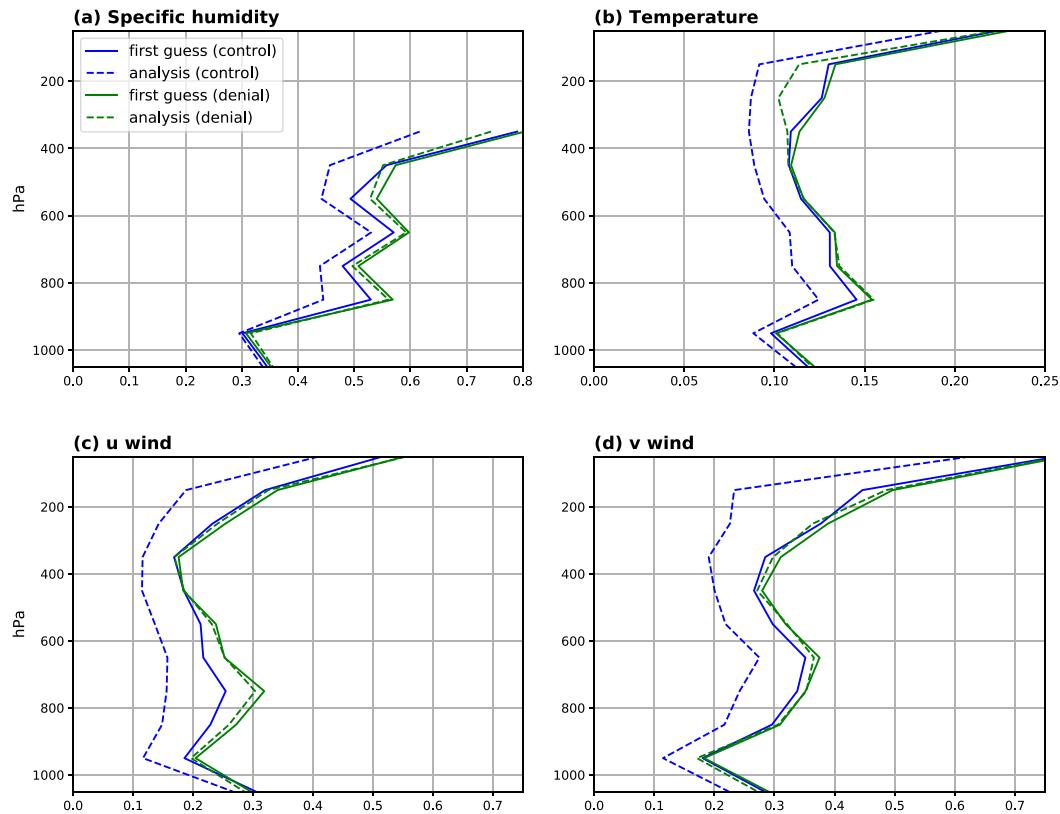


FIG. 8. Normalized vertical profiles of the low-res control RMS differences $[(z_{\text{control}})^{1/2}]$; blue] and the denial RMS differences $[(z_{\text{denial}})^{1/2}]$; green] for the Kiritimati Island and Ron Brown ENRR observations and background (solid) or analysis (dashed) of (a) specific humidity, (b) temperature, (c) zonal wind, and (d) meridional wind, interpolated to the ENRR observation location for all times during 14 Feb–7 Mar.

existing observational network (Figs. 4 and 5), the strengthening of this increment by the research-quality ENRR observations, and the slight (though insignificant) improvement in local precipitation estimates suggest an error in which the model does not produce upper-level outflow in this region that is as strong as reality.

4. Local effects of observations

Results comparing the control and denial experiments demonstrate that the ENRR observations improved the 6-h background fields and significantly impacted the resulting analyses at the ENRR radiosonde observation locations. Let $z_{\text{control}} = \frac{\sum_{t \in \text{time}} [x_{\text{ob}}(t) - x_{\text{model}}(t)]^2}{\sum_{t \in \text{time}} [x_{\text{ob}}(t) - \bar{x}_{\text{ob}}]^2}$ be the mean squared difference (MSD) between the observations x_{ob} and the background or analysis field x_{model} from the control run interpolated to the observation time and location, normalized by the variance of the observations, for $x =$ specific humidity, temperature, zonal wind, and meridional wind, where \bar{x}_{ob} denotes the mean of the observations. Similarly, define z_{denial} as the normalized

mean squared difference between the observations and the fields from the denial experiment. Figure 8 shows the vertical profiles of the root-mean-squared (RMS) differences $(z_{\text{control}})^{1/2}$ (blue) and $(z_{\text{denial}})^{1/2}$ (green) for the background (solid) and analysis (dashed) fields. The mean is taken over ENRR radiosonde platforms (Kiritimati Island and RHB) and in time (14 February–7 March 2016, all hours). While the dashed curves are not independent of the ENRR observations (since the control analysis fields assimilated them), we emphasize that the solid curves are independent of the ENRR observations used within the comparison (since the first-guess fields are compared to the ENRR observations that have not yet been assimilated). Results including ENRR dropsondes, in addition to radiosondes, are qualitatively similar (not shown). Note that this calculation covers a shorter time period than the 22 January–7 March 2016 period used in earlier sections, because the Kiritimati Island observations were not assimilated until after 14 February 2016.

To determine whether the differences between the control and denial experiments are significant, Fig. 9

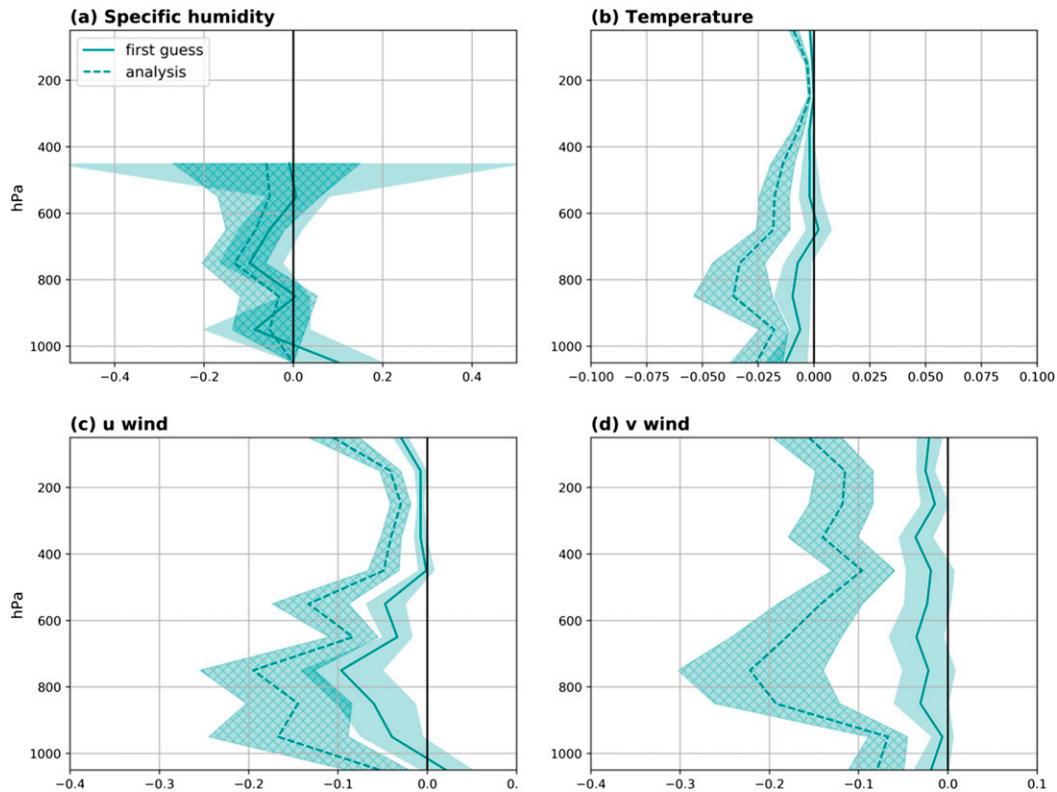


FIG. 9. Normalized vertical profiles of the difference between the low-res control MSD (z_{control}) and the denial MSD (z_{denial}) for the Kiritimati Island and Ron Brown ENRR observations and background (solid) or analysis (dashed) of (a) specific humidity, (b) temperature, (c) zonal wind, and (d) meridional wind interpolated to the ENRR observation location for all times during 14 Feb–7 Mar. Shading represents 95% confidence intervals derived from a paired block bootstrap.

shows the vertical profiles of the differences ($z_{\text{control}} - z_{\text{denial}}$) for the low-res experiment with 95% confidence intervals. Profiles to the left of the zero line demonstrate a positive impact of assimilating ENRR observations. Shading represents 95% confidence intervals determined by a paired block bootstrap technique (Hamill 1999). This bootstrap consists of drawing 1000 samples with replacement from the original set of squared differences between the observation and the interpolated background (analysis) field, pairing the control and denial experiments. An empirical distribution of the sample means is then constructed, from which confidence intervals can be determined. The resampling is conducted in 6-h blocks of time, under the assumption that observations 6 h apart are reasonably independent while observations within the same 6-h window are not assumed to be independent.

It is perhaps unsurprising that the analysis fields are pulled closer to the ENRR observations in the control experiment (when they were assimilated), relative to the denial experiment (when they were not), as shown by the dashed curves in Fig. 9. This implies that the

routine observing system, without ENRR observations, did not produce analyses that were as close to the ENRR observations as they could have been with a more complete observing system. In other words, the ENRR observations were not redundant in the existing observational network. Demonstrating their beneficial impact, assimilating ENRR observations brings the first-guess fields closer to the ENRR observations (solid curves in Fig. 9) for all variables considered at nearly all levels. Exceptions include surface-level specific humidity and zonal wind, but these differences are statistically insignificant. The background fields are consistently, and often significantly, closer to the ENRR radiosonde observations in the low-res control experiment than in the denial experiment. The beneficial impacts of the observations are particularly notable in the middle troposphere: the differences between the control and denial analysis fit to the observations is largest from about 500 to 800 hPa in all variables, where the observation density is low (Figs. 1, 2). Particularly notable in this measure is how much the assimilation of ENRR observations

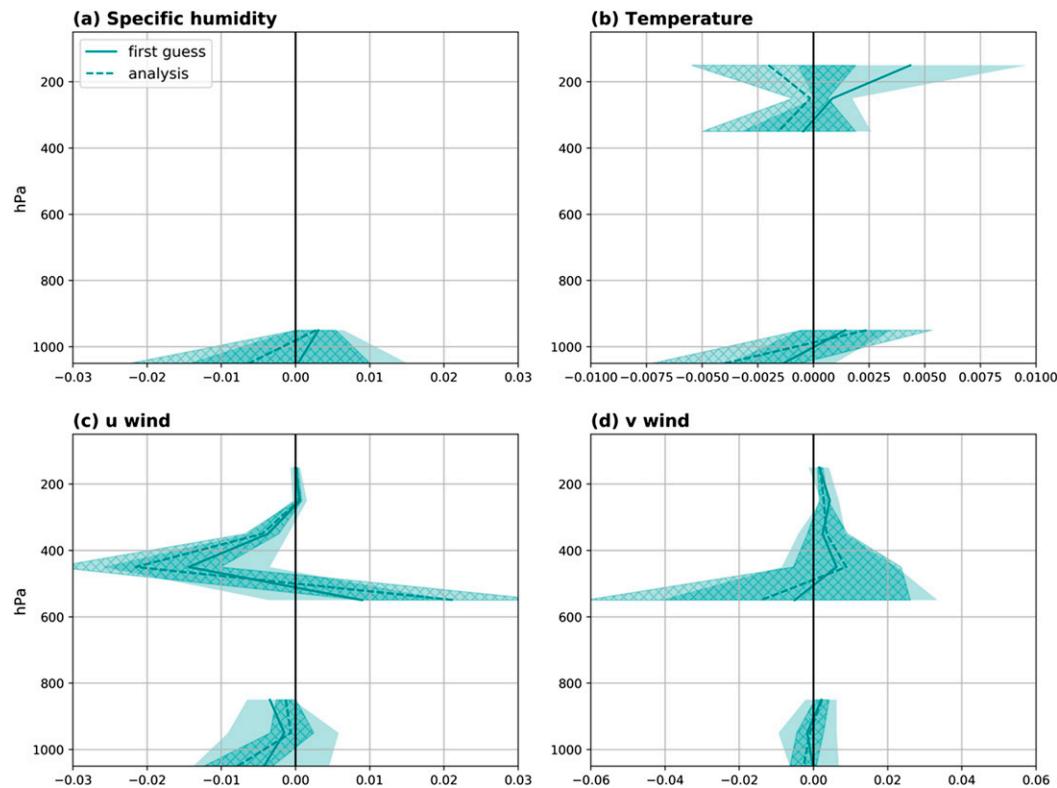


FIG. 10. Normalized vertical profiles of the difference between the low-res control MSD (z_{control}) and the denial MSD (z_{denial}) for non-ENRR observations in region A and background (solid) or analysis (dashed) of (a) specific humidity, (b) temperature, (c) zonal wind, and (d) meridional wind interpolated to the observation location for all times during 14 Feb–7 Mar. Shading represents 95% confidence intervals derived from a paired block bootstrap.

significantly improved zonal wind background fields at these levels.

5. Comparisons to other observations

Figure 5 suggests that the ENRR observations had an impact on the background and analysis fields of 200-hPa meridional wind within a region of the deep tropics. To determine the effect of the ENRR observations on other variables in that region, the analysis and first-guess fields of temperature, humidity, and winds are compared with non-ENRR observations in a region of the deep tropical Pacific Ocean defined by 2°S – 19°N , 180° – 120°W ; see region A of Fig. 3 (magenta box).

Despite the significant beneficial effects that the ENRR observations have on the background and analysis fields interpolated to their observation times and locations (e.g., Fig. 9), the effects are much weaker when assessed using other observations, as illustrated in Fig. 10. These vertical profiles show the normalized difference between the low-res control MSD (z_{control}) and the denial MSD (z_{denial}) assessed using only non-ENRR observations, including in situ observations as

well as aircraft and satellite winds, within this region and that were assimilated in all experiments. First guess and analysis fields have been interpolated to each observation time and location. The MSD's from the control experiment that assimilated the ENRR observations are not consistently smaller than those from the denial experiment, and these differences are generally within the 95% confidence intervals. The exception is 350–450-hPa zonal wind in which the control fields are significantly closer to the observations than the denial fields [negative ($z_{\text{control}} - z_{\text{denial}}$)]. Note also that for 200–300-hPa meridional wind, the control fields are slightly but significantly *farther* from the observations than the denial fields [positive ($z_{\text{control}} - z_{\text{denial}}$)]. It is important to note that all of the differences in Fig. 10 are about 1/10 the magnitude than the differences in Fig. 9, indicating that the analysis and first guess fits to the non-ENRR observations are virtually unchanged by the assimilation of ENRR observations.

Over the eastern tropical Pacific Ocean, there are very few in situ wind observations (e.g., Fig. 1); nearly all of the non-ENRR wind observations in Fig. 10 are from satellites and aircraft. To determine whether

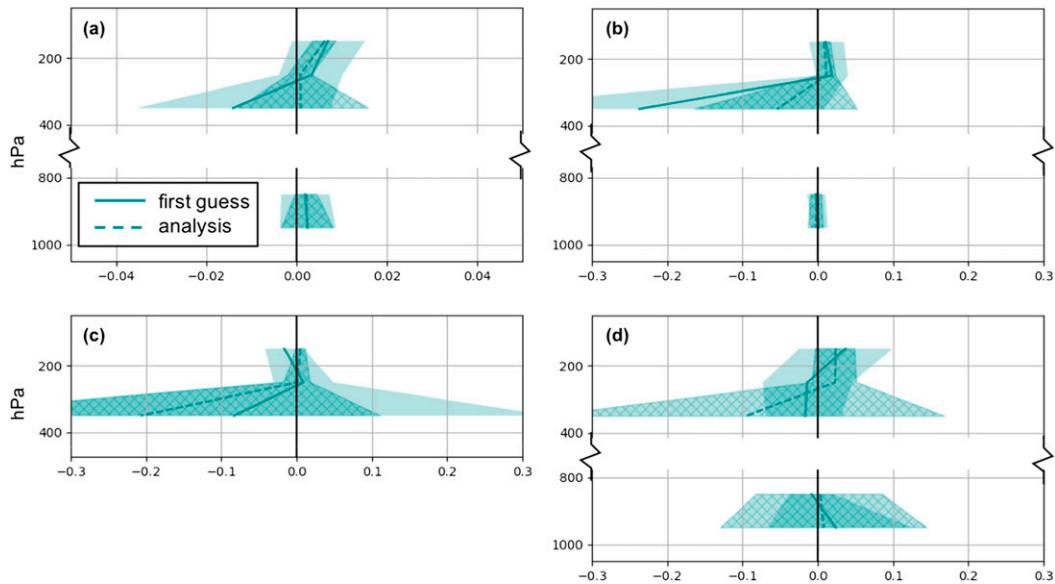


FIG. 11. Normalized vertical profiles of the difference between the low-res control MSD (z_{control}) and the denial MSD (z_{denial}) for AMV v -wind observations (a) within region B for all times during the 14 Feb–7 Mar period; (b) within region B for “deep tropics” flight days, valid at 0000 UTC; (c) within region C for “deep tropics” flight days, valid at 0000 UTC; and (d) within region D for “deep tropics” flight days, valid at 0000 UTC.

assimilating the additional in situ observations from the ENRR campaign improved the fit of the first guess and analysis wind fields specifically to atmospheric motion vector (AMV) observations from satellites, these fields are compared to the AMV observations in three selected regions of the tropical Pacific Ocean. Region B (gold dashed box of Fig. 3) is defined as the region 2°S–19°N by 170°–140°W. The deep tropics region is then split into regions C and D (blue dotted and green dash-dotted boxes of Fig. 3, respectively): 2°S–4°N by 165°–150°W and 4°–10°N by 165°–150°W. Differences are first considered in region B for the longer time period of 14 February–7 March, before narrowing the time period to 0000 UTC on deep tropics flight days, and finally considering regions C and D on deep tropics flight days (Figs. 11 and 12).

Similar to the results of Fig. 10, control and denial differences with AMVs are inconsistent and generally insignificant. In particular, the mean squared differences between observations and fields at 200 hPa are not consistently smaller in the control experiments than in the denial experiments (Fig. 11), suggesting that the large increments in the deep tropics box of Fig. 5 may not be drawing the fields closer to the AMV observations in a mean-square sense.

Investigations into the mean biases $y = \left\{ \frac{\sum_{t \in \text{time}} [x_{\text{ob}}(t) - x_{\text{model}}(t)]}{\sum_{t \in \text{time}} [x_{\text{ob}}(t) - \bar{x}_{\text{ob}}]^2} \right\}^{1/2}$ are similarly inconsistent, with few significant results (Fig. 12). On deep tropics flight days, in region C (Fig. 12c),

the control background fields appear to be closer to the AMV meridional wind observations than the denial fields at all levels, while the control analysis fields are farther away. In region D (Fig. 12d), the results are mixed. Consistent with Fig. 5, results shown in Fig. 12d suggest that both the control background and analysis fields at 200 hPa are closer to the AMV observations in this region. The largely insignificant and varying results when comparing to AMV observations may be an effect of their relatively large observation errors: the AMVs have errors that are roughly twice as large as those from radiosondes and dropsondes (AMV errors vary from 3.8 to 14 m s⁻¹, while errors from in situ platforms vary from 1.4 to 3.4 m s⁻¹), so the analysis field may not be expected to fit the AMV observations as well as it fits the in situ observations. It is therefore difficult to conclude whether the increments in Fig. 5 from the control experiments result in analyses and background fields that are closer to AMV observations.

6. High-res case results

Concurrent with the low-res experiments, a separate set of high-res experiments was also conducted to determine the impact of the ENRR observations on an operational weather forecast system. As discussed earlier, these high-res experiments were run with an ensemble at a resolution of T574 and a control member at a resolution of T1534, using the hybrid 4DEnVar that

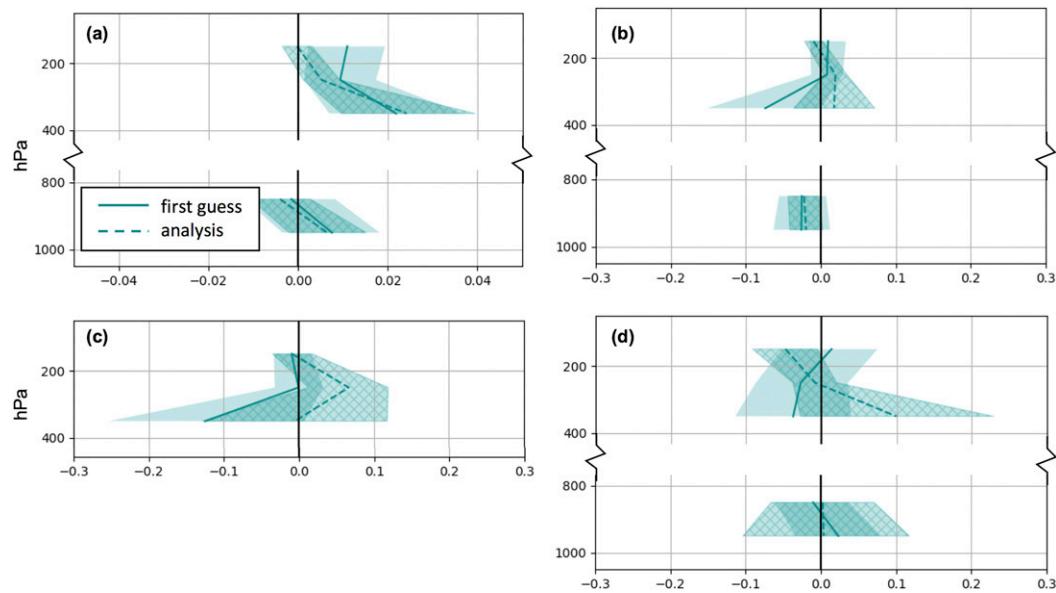


FIG. 12. Normalized vertical profiles of the difference between the low-res control bias (v_{control}) and the denial bias (v_{denial}) for AMV v -wind observations (a) within region B for all times during the 14 Feb–7 Mar period; (b) within region B for “deep tropics” flight days, valid at 0000 UTC; (c) within region C for “deep tropics” flight days, valid at 0000 UTC; and (d) within region D for “deep tropics” flight days, valid at 0000 UTC.

was operational at the Environmental Modeling Center in May 2016. These two sets of experiments (low-res and high-res) are not intended to be directly compared with one another, as there are several differences between them apart from the resolution. First, recall (section 2) that the low-res experiments use a 4DIAU algorithm and do not recenter the ensemble around a control member, while the high-res experiments do not use the 4DIAU algorithm and do recenter the ensemble around a control member. Second, while both sets of ensembles (high-res and low-res) use the SPPT stochastic perturbation on precipitation, the high-res control member does not stochastically perturb precipitation. Third, the low-res experiments assimilated wind observations between 450 and 550 hPa from the deep-layer water vapor imager on the *Geostationary Operational Environmental Satellite-13* (*GOES-13*). These observations were blacklisted operationally, and thus were unused in the high-res experiment. An in-depth investigation involving another set of data-denial experiments would be needed to determine the impact of the *GOES-13* observations; however, these data are no longer produced. They have been replaced by the deep-level water vapor AMVs observed by the advanced baseline imager on *GOES-16*, which have been assimilated operationally since January 2018 (NOAA/EMC 2018). Several other small differences in model configurations exist between the low-res and high-res experiments,

including choice of land surface datasets and ozone parameterizations.

However, the high-res control and high-res denial experiments only differed in their assimilation or denial of the ENRR observations, as in the low-res experimental setup. Therefore, the results from the high-res experiments can be used to support the conclusions drawn from the low-res experiments above, regarding the impact of the ENRR observations. The high-res 200-hPa meridional wind fields are similar to those from the respective low-res experiments (see Figs. 13 and 14). In the high-res experiments, the difference between the control and denial increments on deep tropics flight days is more localized to the tropics than in the low-res experiments (cf. Fig. 14c with Fig. 5c), and has less statistical significance: the low-res differences in Fig. 5c are significant at the 90% level (not shown), but the high-res differences in Fig. 14c have no areas of statistical significance at the 90% level.

The local effects of the ENRR observations on the high-res experiments are also similar to the low-res results: the control MSD is smaller than the denial MSD for most variables at most levels, though there is overall less significance in these differences (Fig. 15). These results suggest that an increased model resolution may allow the assimilation system to use the existing observational network more effectively than the low-res case. However, this could also be a result of the IAU: if the high-res increments (without IAU) are larger than the

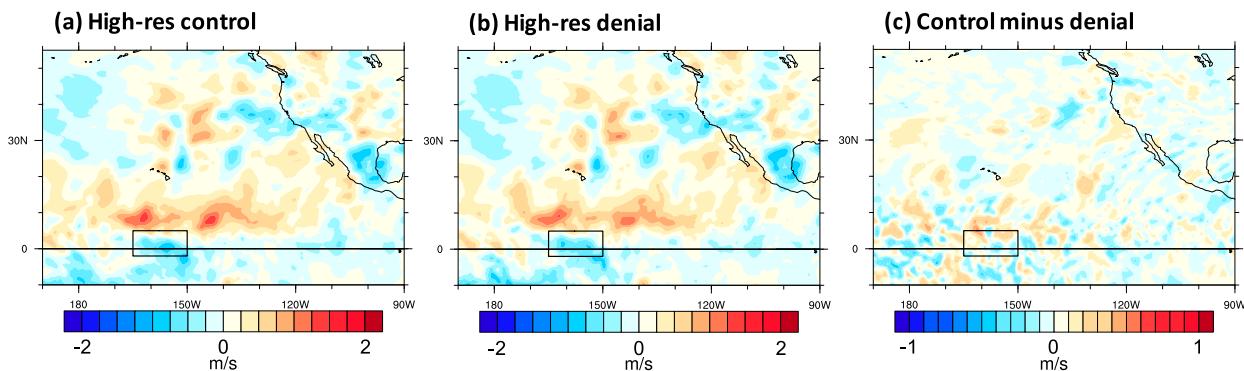


FIG. 13. Increments of 200-hPa meridional wind, as in Fig. 4, but for the high-res experiment.

low-res increments post-IAU, this could lead to subsequent forecasts from the high-res analysis fields being closer to the observations. Because of the inconsistencies in the two experimental designs, however, these differences cannot for certain be attributed to the model resolution or to the IAU, and further work is required to rigorously determine the effects of each of these on the GFS assimilation system. Comparisons to nearby non-ENRR radiosonde observations and satellite wind data are similar in the high-res case (not shown) to the low-res case (Figs. 9–11): differences between the high-res control and denial are inconsistent across variable and vertical level.

7. Conclusions

Results from two sets of data-denial experiments demonstrate the impacts of the ENRR field campaign observations on the analysis and 6-h background fields of temperature, specific humidity, and wind. First, assimilating ENRR observations led to stronger outflow associated with 200-hPa divergence from the deep tropics region of the Pacific Ocean, particularly when flights entered that area. The systematic increments of 200-hPa meridional wind seem to point to a model bias

in this region, as the background fields consistently showed weaker outflow than the observations demonstrated. Wang et al. (2019) suggest that this systematic strengthening of upper-level outflow acts to strengthen El Niño-related features in 7-day forecasts (their Fig. 9b). This signal is consistent with earlier studies of the effects of anomalous equatorial Pacific heat sources (Ting and Sardeshmukh 1993).

Second, assimilating ENRR observations pulled the analysis fields closer to the observations. While perhaps unsurprising, this does suggest that these observations were not redundant within the existing observation network, despite the dense satellite and aircraft coverage of some variables throughout many levels of the troposphere. Notably, the existing wind observation network has gaps in the middle troposphere, where the observations significantly affected the analyses.

Third, 6-h background fields show small but consistent and significant differences in temperature, specific humidity, and wind, suggesting that the ENRR observations improved the background fields locally. This is consistent with the results of Wang et al. (2019), who show significant impacts of the ENRR observations on global, tropical, and hemispheric measures of 12–24 h forecast skill in the same NCEP GFS model used here.

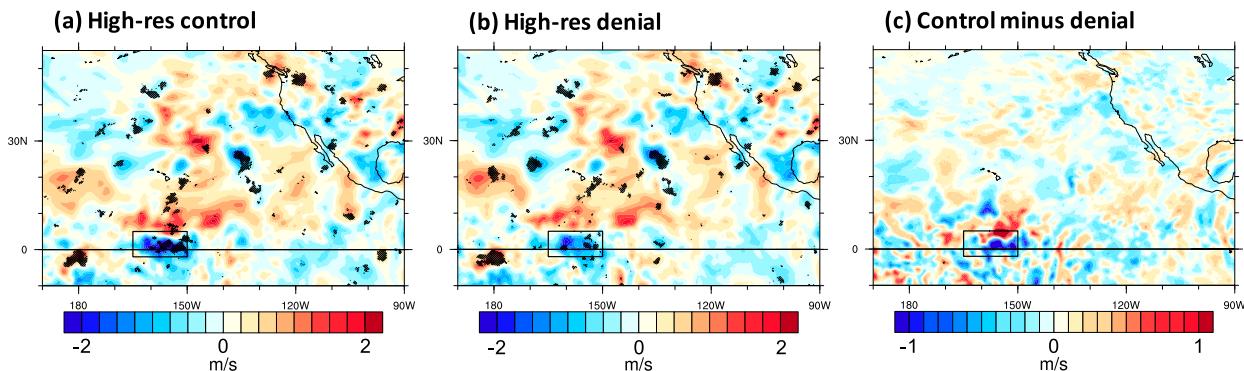


FIG. 14. Increments of 200-hPa meridional wind, as in Fig. 5, but for the high-res experiment.

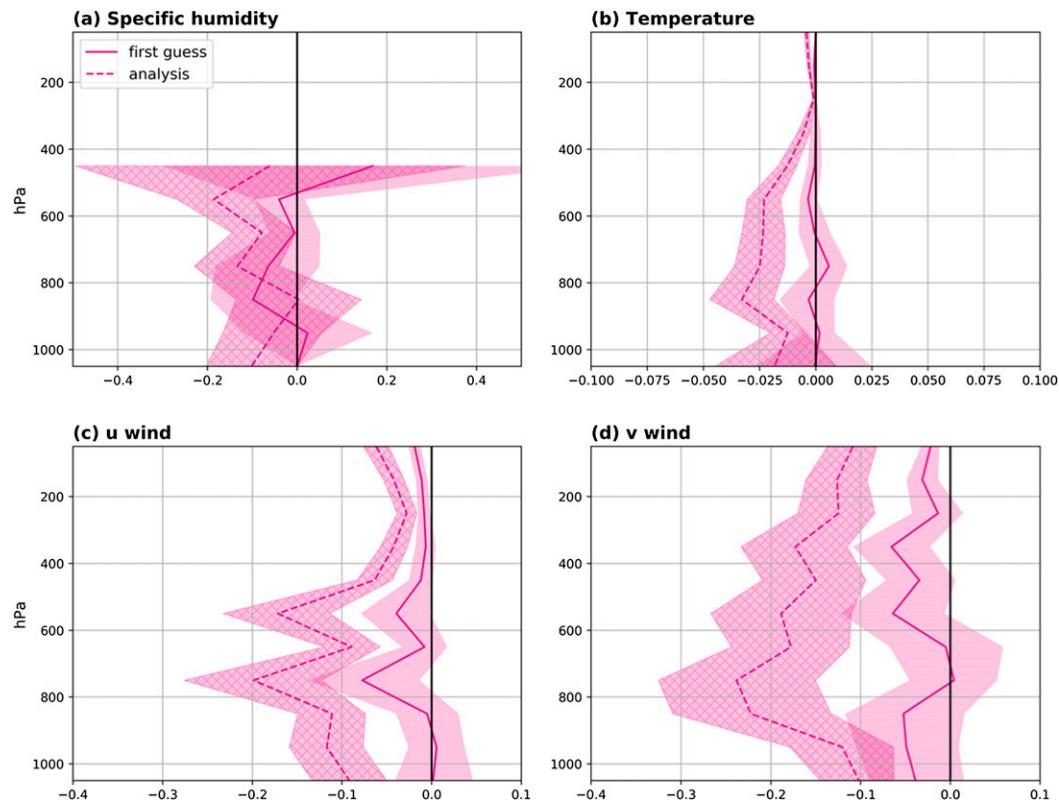


FIG. 15. Normalized vertical difference profiles, as in Fig. 9, but for the high-res experiments.

This suggests that the localized impacts on the analysis and 6-h background fields are communicated to those regions over time. However, the authors also found that the impacts of these observations are negligible past forecast hour 24 (their Fig. 2), suggesting that the small but significant impacts of the ENRR observations at up to 24-h leads are lost for longer leads.

Despite these interesting results, the ENRR observations had mixed effects on the fit of the analysis or background fields to nearby non-ENRR observations. The strong increments of 200-hPa meridional wind caused by the ENRR observations correspond to analyses that are closer (in a mean sense) to AMV observations in a small region of the tropical Pacific Ocean, suggesting that assimilating the ENRR observations helped to correct a model bias in this region. However, this result is not robust across different time periods and spatial regions, nor the statistic used, since the mean-squared differences do not show consistent improvement when the ENRR observations are assimilated. Effects of the additional ENRR observations on precipitation fields are similarly insignificant.

Finally, results from a similar set of data-denial experiments that used a higher-resolution forecast model

from an operational system support the conclusions above. Despite several other differences in experimental design, the addition of ENRR observations within the high-res system strengthens the increment of 200-hPa divergence, improves the fit of background and analysis fields to the ENRR observations, and has mixed effects on the fit to nearby non-ENRR observations, comparable to the low-res experiments.

Overall, the ENRR campaign provided many additional useful observations over the tropical Pacific Ocean during a major El Niño event. These observations had significant local effects on the output of the assimilation system, but large-scale improvements from them were limited. While additional independent observations can be helpful for validation, we suggest that future observational field campaigns would benefit from detailed prior knowledge of the impact of different types of observations within the existing network, particularly of satellite and aircraft observations. Other data-denial experiments, in which the satellite or aircraft observing networks are systematically degraded within an operational forecast and assimilation framework, would likely shed more light on the impact of these observations. A suite of such data-denial experiments would reveal the relative impact of different types of observations, and

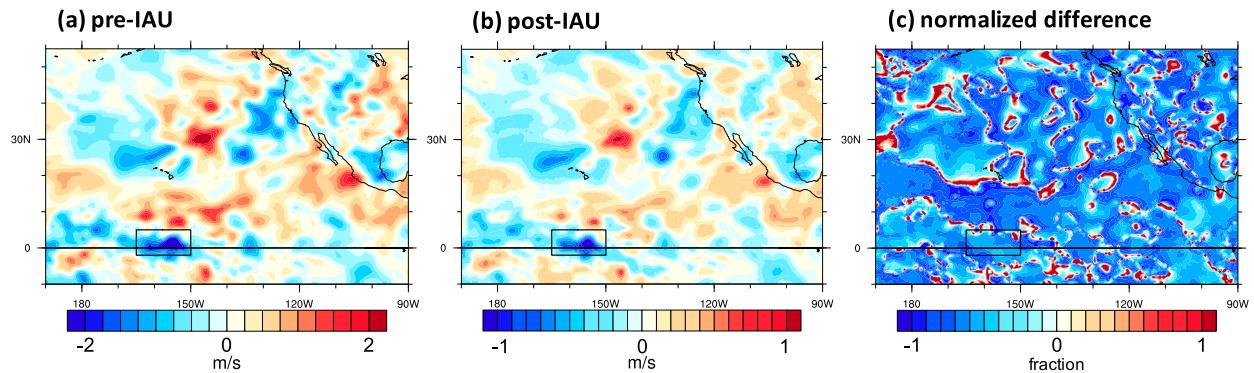


FIG. A1. Average 200-hPa meridional wind increments calculated using (a) the initial pre-IAU analysis; (b) the final post-IAU analysis; and (c) the normalized difference of the magnitude of the two (see text for details). The average is taken over deep tropics flight days, 0000 UTC. Note the different color scales used for (a) and (b).

provide a framework for determining where future efforts should be focused.

Acknowledgments. We thank three anonymous reviewers, as well as R.S. Webb and T. Hamill of NOAA Earth System Research Laboratory's Physical Sciences Division, for helpful comments on this work. We greatly appreciate the contributions from many people within NOAA (OAR, NWS, OMAO, NESDIS) who assisted in the collection of data as part of the El Niño Rapid Response Field Campaign. NOAA OAR Physical Sciences Division resources supported the modeling and analysis research. Support was also provided by the NOAA Climate Program Office. Computing was performed on NOAA's Remotely Deployed High Performance Computing System Theia and the IBM WCOSS platform. The scientific results and conclusions, as well as any views or opinions expressed herein, are those of the authors and do not necessarily reflect the views of NOAA or the Department of Commerce.

APPENDIX

Effect of Incremental Analysis Updates

The low-res experiments in this work use a 4D incremental analysis update (4DIAU; Bloom et al. 1996; Lei and Whitaker 2016); see section 2. Briefly, the traditional or "3DIAU" consists of applying each DA increment as a constant forcing within the model, between assimilation times. Since its introduction by Bloom et al. (1996), this method has been widely implemented in many different assimilation systems, with demonstrations of its effects including damped unstable subspaces, decreased tendencies, reduced discontinuities, and smaller increments (Zhu et al. 2003; Polavarapu et al. 2004; Zhang et al. 2015; Ha et al. 2017).

The 4DIAU, first discussed by Lorenc et al. (2015), differs from the 3DIAU in that the increments are calculated and applied over smaller windows of time within each assimilation time step, whereas the 3DIAU calculates one increment for each assimilation window. The 4DIAU was implemented successfully within the 4DEnVar in the U.S. Global Forecast System, with improvements over 3DIAU (Lei and Whitaker 2016). It was also implemented within the 4DEnVar at Environment and Climate Change Canada, where it was shown to reduce spurious wave activity generated by imbalances as compared to digital filtering (Buehner et al. 2015).

In the low-res ENRR experiments, there is a difference in magnitude between the 200-hPa meridional wind fields from the analysis increment (analysis minus forecast) calculated before applying IAU and from the increment calculated after the IAU was applied (valid at the same time). Figure A1 shows the increment of 200-hPa meridional wind, averaged over the 10 deep tropics flight days (0000 UTC), calculated using the pre-IAU analysis (Fig. A1a) and the post-IAU analysis (Fig. A1b). Note that the IAU window spans 6 h from (analysis time minus 3 h) to (analysis time plus 3 h), so the post-IAU fields at 0000 UTC have had 3 h of IAU forcing, not the full 6 h. Figure A1c shows the difference between the magnitudes (absolute value) of the post-IAU and pre-IAU analysis increments, normalized by the pre-IAU increment magnitude. The difference plot demonstrates that the post-IAU increment is often 60%–80% weaker than the pre-IAU increment. (The dark red areas correspond to regions where both increments are nearly 0.) Note that differences may be especially large for a field like 200-hPa meridional wind and other fields related to divergence, where the instantaneous increment may be inconsistent with the model's hydrologic cycle in the tropics. Improvements in the model's physics could

result in a decrease in this mis-match; investigations into this question are outside the scope of this paper.

Future work is planned to investigate the effects of the 4DIAU in more detail. Questions include: how the IAU affects forecasts on longer time scales (e.g., 6–12 h after the IAU forcing is turned off); the effect of the IAU on other variables; whether and how differences between pre and post IAU fields can be used to diagnose model error; and average differences between increments from digital filtering (without IAU), from the final IAU fields, and from the pre-IAU fields.

REFERENCES

- Anwender, D., C. Cardinali, and S. C. Jones, 2012: Data denial experiments for extratropical transition. *Tellus*, **64A**, 19151, <https://doi.org/10.3402/tellusa.v64i0.19151>.
- Barsugli, J. J., and P. D. Sardeshmukh, 2002: Global atmospheric sensitivity to tropical SST anomalies throughout the Indo-Pacific basin. *J. Climate*, **15**, 3427–3442, [https://doi.org/10.1175/1520-0442\(2002\)015<3427:GASTTS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<3427:GASTTS>2.0.CO;2).
- Benjamin, S. G., B. E. Schwartz, S. E. Koch, and E. J. Szoke, 2004: The value of wind profiler data in U.S. weather forecasting. *Bull. Amer. Meteor. Soc.*, **85**, 1871–1886, <https://doi.org/10.1175/BAMS-85-12-1871>.
- Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271, [https://doi.org/10.1175/1520-0493\(1996\)124<1256:DAUIAU>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2).
- Bormann, N., S. Saarinen, G. Kelly, and J.-N. Thépaut, 2003: The spatial structure of observation errors in atmospheric motion vectors from geostationary satellite data. *Mon. Wea. Rev.*, **131**, 706–718, [https://doi.org/10.1175/1520-0493\(2003\)131<0706:TSSOOE>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0706:TSSOOE>2.0.CO;2).
- Bouttier, F., and G. Kelly, 2001: Observing-system experiments in the ECMWF 4D-Var data assimilation system. *Quart. J. Roy. Meteor. Soc.*, **127**, 1469–1488, <https://doi.org/10.1002/qj.49712757419>.
- Buehner, M., and Coauthors, 2015: Implementation of deterministic weather forecasting systems based on ensemble-variational data assimilation at Environment Canada. Part I: The global system. *Mon. Wea. Rev.*, **143**, 2532–2559, <https://doi.org/10.1175/MWR-D-14-00354.1>.
- Compo, G. P., P. D. Sardeshmukh, and C. Penland, 2001: Changes of subseasonal variability associated with El Niño. *J. Climate*, **14**, 3356–3374, [https://doi.org/10.1175/1520-0442\(2001\)014<3356:COVAV>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<3356:COVAV>2.0.CO;2).
- Cox, C. J., D. E. Wolfe, L. M. Hartten, and P. E. Johnston, 2017: El Niño Rapid Response (ENRR) Field Campaign: Radiosonde Data (Level 2) from the NOAA Ship Ronald H. Brown, February–March 2016 (NCEI Accession 0161527). NOAA National Centers for Environmental Information, accessed 1 November 2016, <https://doi.org/10.7289/v5x63k15>.
- Dole, R., and Coauthors, 2018: Advancing science and services during the 2015/16 El Niño: The NOAA El Niño Rapid Response Field Campaign. *Bull. Amer. Meteor. Soc.*, **99**, 975–1001, <https://doi.org/10.1175/BAMS-D-16-0219.1>.
- Gelaro, R., and Y. Zhu, 2009: Examination of observation impacts derived from observing system experiments (OSEs) and adjoint models. *Tellus*, **61A**, 179–193, <https://doi.org/10.1111/j.1600-0870.2008.00388.x>.
- , R. H. Langland, S. Pellerin, and R. Todling, 2010: The THORPEX observation impact intercomparison experiment. *Mon. Wea. Rev.*, **138**, 4009–4025, <https://doi.org/10.1175/2010MWR3393.1>.
- Glantz, M. H., 2001: *Currents of Change: Impacts of El Niño and La Niña on Climate and Society*. 2nd ed. Cambridge University Press, 268 pp.
- Ha, S., C. Snyder, W. C. Skamarock, J. Anderson, and N. Collins, 2017: Ensemble Kalman filter data assimilation for the Model for Prediction Across Scales (MPAS). *Mon. Wea. Rev.*, **145**, 4673–4692, <https://doi.org/10.1175/2017-0145.1>.
- Halpert, M. S., and C. F. Ropelewski, 1992: Surface temperature patterns associated with the Southern Oscillation. *J. Climate*, **5**, 577–593, [https://doi.org/10.1175/1520-0442\(1992\)005<0577:STPAWT>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<0577:STPAWT>2.0.CO;2).
- Hamill, T. M., 1999: Hypothesis tests for evaluating numerical precipitation forecasts. *Wea. Forecasting*, **14**, 155–167, [https://doi.org/10.1175/1520-0434\(1999\)014<0155:HTFENP>2.0.CO;2](https://doi.org/10.1175/1520-0434(1999)014<0155:HTFENP>2.0.CO;2).
- , F. Yang, C. Cardinali, and S. J. Majumdar, 2013: Impact of targeted winter storm reconnaissance dropwindsonde data on midlatitude numerical weather predictions. *Mon. Wea. Rev.*, **141**, 2058–2065, <https://doi.org/10.1175/MWR-D-12-00309.1>.
- Harnisch, F., and M. Weissmann, 2010: Sensitivity of typhoon forecasts to different subsets of targeted dropsonde observations. *Mon. Wea. Rev.*, **138**, 2664–2680, <https://doi.org/10.1175/2010MWR3309.1>.
- Hartten, L. M., P. E. Johnston, C. J. Cox, and D. E. Wolfe, 2017: El Niño Rapid Response (ENRR) Field Campaign: Surface meteorological data from Kiritimati Island, January–March 2016 (NCEI Accession 0161526). NOAA National Centers for Environmental Information, accessed 1 November 2016, <https://doi.org/10.7289/v51z42h4>.
- , C. J. Cox, P. E. Johnston, D. E. Wolfe, S. Abbott, and H. A. McColl, 2018a: Central-Pacific surface meteorology from the 2016 El Niño Rapid Response (ENRR) field campaign. *Earth Syst. Sci. Data*, **10**, 1139, <https://doi.org/10.5194/essd-10-1139-2018>.
- , —, —, —, —, —, X.-W. Quan, and M. G. Winterkorn, 2018b: Ship- and island-based soundings from the 2016 El Niño Rapid Response (ENRR) field campaign. *Earth Syst. Sci. Data*, **10**, 1165–1183, <https://doi.org/10.5194/essd-10-1165-2018>.
- Hu, M., C. Zhou, H. Shao, D. Stark, and K. Newman, 2016: Gridpoint Statistical Interpolation (GSI) advanced user's guide version 3.5. Developmental Testbed Center, 124 pp., https://dtcenter.org/com-GSI/users/docs/users_guide/AdvancedGSIUserGuide_v3.5.0.0.pdf.
- Huang, X.-Y., and P. Lynch, 1993: Diabatic digital-filtering initialization: Application to the HIRLAM model. *Mon. Wea. Rev.*, **121**, 589–603, [https://doi.org/10.1175/1520-0493\(1993\)121<0589:DDFIAT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1993)121<0589:DDFIAT>2.0.CO;2).
- Kashino, Y., N. España, F. Syamsudin, K. J. Richards, T. Jensen, P. Dutrieux, and A. Ishida, 2009: Observations of the North Equatorial Current, Mindanao Current, and Kuroshio Current system during the 2006/07 El Niño and 2007/08 La Niña. *J. Oceanogr.*, **65**, 325–333, <https://doi.org/10.1007/s10872-009-0030-z>.
- Kelly, G., J.-N. Thépaut, R. Buizza, and C. Cardinali, 2007: The value of observations. I: Data denial experiments for the Atlantic and the Pacific. *Quart. J. Roy. Meteor. Soc.*, **133**, 1803–1815, <https://doi.org/10.1002/qj.150>.
- Kiladis, G. N., and H. F. Diaz, 1989: Global climatic anomalies associated with extremes in the Southern Oscillation. *J. Climate*,

- 2, 1069–1090, [https://doi.org/10.1175/1520-0442\(1989\)002<1069:GCAAWE>2.0.CO;2](https://doi.org/10.1175/1520-0442(1989)002<1069:GCAAWE>2.0.CO;2).
- Kleist, D. T., and K. Ide, 2015: An OSSE-based evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS. Part II: 4DVar and hybrid variants. *Mon. Wea. Rev.*, **143**, 452–470, <https://doi.org/10.1175/MWR-D-13-00350.1>.
- Langland, R. H., 2005: Issues in targeted observing. *Quart. J. Roy. Meteor. Soc.*, **131**, 3409–3425, <https://doi.org/10.1256/qj.05.130>.
- Lei, L., and J. S. Whitaker, 2016: A four-dimensional incremental analysis update for the ensemble Kalman filter. *Mon. Wea. Rev.*, **144**, 2605–2621, <https://doi.org/10.1175/MWR-D-15-0246.1>.
- L'Heureux, M. L., and Coauthors, 2017: Observing and predicting the 2015/16 El Niño. *Bull. Amer. Meteor. Soc.*, **98**, 1363–1382, <https://doi.org/10.1175/BAMS-D-16-0009.1>.
- Lorenc, A. C., N. E. Bowler, A. M. Clayton, S. R. Pring, and D. Fairbairn, 2015: Comparison of hybrid-4DVar and hybrid-4DVar data assimilation methods for global NWP. *Mon. Wea. Rev.*, **143**, 212–229, <https://doi.org/10.1175/MWR-D-14-00195.1>.
- Lynch, P., and X.-Y. Huang, 1992: Initialization of the HIRLAM model using a digital filter. *Mon. Wea. Rev.*, **120**, 1019–1034, [https://doi.org/10.1175/1520-0493\(1992\)120<1019:IOTHMU>2.0.CO;2](https://doi.org/10.1175/1520-0493(1992)120<1019:IOTHMU>2.0.CO;2).
- Majumdar, S. J., 2016: A review of targeted observations. *Bull. Amer. Meteor. Soc.*, **97**, 2287–2303, <https://doi.org/10.1175/BAMS-D-14-00259.1>.
- , M. J. Brennan, and K. Howard, 2013: The impact of dropwindsonde and supplemental rawinsonde observations on track forecasts for Hurricane Irene (2011). *Wea. Forecasting*, **28**, 1385–1403, <https://doi.org/10.1175/WAF-D-13-00018.1>.
- Massey, F. J., Jr., 1951: The Kolmogorov–Smirnov test for goodness of fit. *J. Amer. Stat. Assoc.*, **46**, 68–78, <https://doi.org/10.1080/01621459.1951.10500769>.
- McPhaden, M. J., S. E. Zebiak, and M. H. Glantz, 2006: ENSO as an integrating concept in earth science. *Science*, **314**, 1740–1745, <https://doi.org/10.1126/science.1132588>.
- Neelin, J. D., D. S. Battisti, A. C. Hirst, F.-F. Jin, Y. Wakata, T. Yamagata, and S. E. Zebiak, 1998: ENSO theory. *J. Geophys. Res.*, **103**, 14 261–14 290, <https://doi.org/10.1029/97JC03424>.
- NOAA, 2018: NOAA G-IV aircraft. NOAA, accessed 30 March 2018, https://www.esrl.noaa.gov/psd/enso/rapid_response/data_pub/.
- NOAA/EMC, 2018: Satellite historical documentation. NOAA, accessed 30 March 2018, http://www.emc.ncep.noaa.gov/mmb/data_processing/Satellite_Historical_Documentation.htm.
- Palmer, T., R. Buizza, F. Doblas-Reyes, T. Jung, M. Leutbecher, G. Shutts, M. Steinheimer, and A. Weisheimer, 2009: Stochastic parametrization and model uncertainty. ECMWF Tech. Memo. 598, 42 pp., <https://www.ecmwf.int/en/elibrary/11577-stochastic-parametrization-and-model-uncertainty>.
- Polavarapu, S., S. Ren, A. M. Clayton, D. Sankey, and Y. Rochon, 2004: On the relationship between incremental analysis updating and incremental digital filtering. *Mon. Wea. Rev.*, **132**, 2495–2502, [https://doi.org/10.1175/1520-0493\(2004\)132<2495:OTRBA>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<2495:OTRBA>2.0.CO;2).
- Romine, G. S., C. S. Schwartz, R. D. Torn, and M. L. Weisman, 2016: Impact of assimilating dropsonde observations from MPEX on ensemble forecasts of severe weather events. *Mon. Wea. Rev.*, **144**, 3799–3823, <https://doi.org/10.1175/MWR-D-15-0407.1>.
- Ropelewski, C. F., and M. S. Halpert, 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation. *Mon. Wea. Rev.*, **115**, 1606–1626, [https://doi.org/10.1175/1520-0493\(1987\)115<1606:GARSPP>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<1606:GARSPP>2.0.CO;2).
- Sardeshmukh, P. D., and B. J. Hoskins, 1988: The generation of global rotational flow by steady idealized tropical divergence. *J. Atmos. Sci.*, **45**, 1228–1251, [https://doi.org/10.1175/1520-0469\(1988\)045<1228:TGOGRF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<1228:TGOGRF>2.0.CO;2).
- , G. P. Compo, and C. Penland, 2000: Changes of probability associated with El Niño. *J. Climate*, **13**, 4268–4286, [https://doi.org/10.1175/1520-0442\(2000\)013<4268:COPAWE>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<4268:COPAWE>2.0.CO;2).
- Ting, M., and P. D. Sardeshmukh, 1993: Factors determining the extratropical response to equatorial diabatic heating anomalies. *J. Atmos. Sci.*, **50**, 907–918, [https://doi.org/10.1175/1520-0469\(1993\)050<0907:FDTERT>2.0.CO;2](https://doi.org/10.1175/1520-0469(1993)050<0907:FDTERT>2.0.CO;2).
- Tompkins, A., and J. Berner, 2008: A stochastic convective approach to account for model uncertainty due to unresolved humidity variability. *J. Geophys. Res.*, **113**, D18101, <https://doi.org/10.1029/2007JD009284>.
- Torn, R. D., 2014: The impact of targeted dropwindsonde observations on tropical cyclone intensity forecasts of four weak systems during PREDICT. *Mon. Wea. Rev.*, **142**, 2860–2878, <https://doi.org/10.1175/MWR-D-13-00284.1>.
- Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998: Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures. *J. Geophys. Res.*, **103**, 14 291–14 324, <https://doi.org/10.1029/97JC01444>.
- UCAR/NCAR–Earth Observing Laboratory, 1994: NSF/NCAR Hercules C130 Aircraft. UCAR/NCAR, accessed 1 November 2016, <https://doi.org/10.5065/D6WM1BG0>.
- Velden, C. S., C. M. Hayden, S. J. Nieman, W. P. Menzel, S. Wanzong, and J. S. Goerss, 1997: Upper-tropospheric winds derived from geostationary satellite water vapor observations. *Bull. Amer. Meteor. Soc.*, **78**, 173–195, [https://doi.org/10.1175/1520-0477\(1997\)078<0173:UTWDFG>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<0173:UTWDFG>2.0.CO;2).
- Wang, J.-W. A., P. D. Sardeshmukh, G. P. Compo, J. S. Whitaker, L. C. Slivinski, C. McColl, and P. Pegion, 2019: Sensitivities of the NCEP Global Forecast System. *Mon. Wea. Rev.*, **147**, 1237–1256, <https://doi.org/10.1175/MWR-D-18-0239.1>.
- Wang, X., D. Parrish, D. Kleist, and J. Whitaker, 2013: GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments. *Mon. Wea. Rev.*, **141**, 4098–4117, <https://doi.org/10.1175/MWR-D-12-00141.1>.
- Zhang, B., V. Tallapragada, F. Weng, J. Sippel, and Z. Ma, 2015: Use of incremental analysis updates in 4D-Var data assimilation. *Adv. Atmos. Sci.*, **32**, 1575–1582, <https://doi.org/10.1007/s00376-015-5041-7>.
- Zhu, Y., R. Todling, J. Guo, S. E. Cohn, I. M. Navon, and Y. Yang, 2003: The GEOS-3 retrospective data assimilation system: The 6-hour lag case. *Mon. Wea. Rev.*, **131**, 2129–2150, [https://doi.org/10.1175/1520-0493\(2003\)131<2129:TGRDAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<2129:TGRDAS>2.0.CO;2).